REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	9/20/01	

4. TITLE AND SUBTITLE
Enzymatic Autocatalysis of Botulinum A Neurotoxin Light Chain

5. FUNDING NUMBERS

6. AUTHOR(S)
S. Ashraf Ahmed, Michael P. Byrne, Melody Jensen, Harry B. Hines, Ernst Brueggemann, and Leonard A. Smith

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
Commander, U.S. Army Medical Research Institute of Infectious Diseases
1425 Porter Street
Fort Detrick, MD 21702-5011

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT
APPROVED FOR PUBLIC RELEASE
DISTRIBUTION UNLIMITED

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Highly purified recombinant zinc-endopeptidase light chain of the botulinum neurotoxin serotype A underwent autocatalytic proteolytic processing and fragmentation. In the absence of added zinc, initially 10-28 residues were cleaved from the C-terminal end of the 448-residue protein followed by the appearance of an SDS-stable dimer and finally fragmentation near the middle of the molecule. In the presence of added zinc, the rate of fragmentation was accelerated but the specificity of the cleavavle bond changed, suggesting a structural role for zinc in the light chain. The C-terminal proteolytic processing was reduced, and fragmentation near the middle of the molecule was prevented by adding the metal chelator TPEN to the light chain. Similarly, adding a competitive peptide inhibitor (CRATKML) of the light-chain catalytic activity also greatly reduced the proteolysis. With these results, for the first time, we provide clear evidence that the loss of C-terminal peptides and fragmentation of the light chain are enzymatic and autocatalytic. By isolating both the large and small peptides, we sequenced them by Edman degradation and ESIMS-MS, and mapped the sites of proteolysis. We also found that proteolysis. We also found that proteolysis occurred at F266-G267, F419-T420, F423-E424, R432-G433, and C430-V431 bonds in addition to the previously reported Y250-Y251 and K438-T439 bonds.

14. SUBJECT TERMS
Botulinum neurotoxin; autocatalysis; proteolytic degradation; zinc-endopeptidase; light chain; synthetic gene

15. NUMBER OF PAGES
11

16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
UNCLASSIFIED	UNCLASSIFIED	UNCLASSIFIED	

Journal of Protein Chemistry, Vol. 20, No. 3, 2001

Enzymatic Autocatalysis of Botulinum A Neurotoxin Light Chain

S. Ashraf Ahmed,[1,4] Michael P. Byrne,[1,3] Melody Jensen,[1] Harry B. Hines,[2] Ernst Brueggemann,[2] and Leonard A. Smith[1]

Received April 19, 2001

Highly purified recombinant zinc-endopeptidase light chain of the botulinum neurotoxin serotype A underwent autocatalytic proteolytic processing and fragmentation. In the absence of added zinc, initially 10–28 residues were cleaved from the C-terminal end of the 448-residue protein followed by the appearance of an SDS-stable dimer and finally fragmentation near the middle of the molecule. In the presence of added zinc, the rate of fragmentation was accelerated but the specificity of the cleavable bond changed, suggesting a structural role for zinc in the light chain. The C-terminal proteolytic processing was reduced, and fragmentation near the middle of the molecule was prevented by adding the metal chelator TPEN to the light chain. Similarly, adding a competitive peptide inhibitor (CRATKML) of the light-chain catalytic activity also greatly reduced the proteolysis. With these results, for the first time, we provide clear evidence that the loss of C-terminal peptides and fragmentation of the light chain are enzymatic and autocatalytic. By isolating both the large and small peptides, we sequenced them by Edman degradation and ESIMS-MS, and mapped the sites of proteolysis. We also found that proteolysis occurred at F266–G267, F419–T420, F423–E424, R432–G433, and C430–V431 bonds in addition to the previously reported Y250–Y251 and K438–T439 bonds.

KEY WORDS: Botulinum neurotoxin; autocatalysis; proteolytic degradation; zinc-endopeptidase; light chain; synthetic gene.

1. INTRODUCTION

Proteolytic modification of proteins often by autocatalytic action plays an essential role in making biologically functional, mature products (Strasser *et al.,* 2000; Dalbey and Kuhn, 2000). Due to the ubiquitous presence of proteases, differentiating endogenous, contaminating protease and autocatalytic protease action on target proteins is difficult. In this paper we address the question: Is C-terminal processing and fragmentation of the light chain (LC[5]) of botulinum neurotoxin serotype A (BoNT/A) autocatalytic?

Botulinum neurotoxins (BoNT), of which seven immunologically distinct serotypes are known, are the most potent of all toxins (For reviews see Montecucco and Schiavo, 1995; Schiavo *et al.,* 1995) These neurotoxins

[1] Department of Immunology and Molecular Biology, U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, Maryland 21702-5011.

[2] Department of Cell Biology and Biochemistry, Division of Toxinology and Aerobiology, U.S. Army Medical Research Institute of Infectious Diseases, Fort Detrick, Maryland 21702-5011.

[3] Present address: Analytical Development, Human Genome Sciences, Rockville, Maryland 20850.

[4] To whom correspondence should be addressed at e-mail: syed.ahmed@amedd.army.mil

[5] Abbreviations: BoNT, botulinum neurotoxin; BoNT/A, botulinum neurotoxin serotype A; LC, light chain; HC, heavy chain; H_n, N-terminal domain of the heavy chain; H_c, C-terminal domain of the heavy chain; SNAP-25, synaptosomal-associated protein of 25 kDa; VAMP, vesicle-associated membrane protein; SNARE, soluble NSF attachment protein receptor; IPTG, isopropyl β-D-thiogalactopyranoside; EDTA, ethylene diamine tetraacetate; TPEN, tetrakis(2-pyridylmethyl)ethylenediamine; SDS–PAGE, sodium dodecyl sulfate–polyacrylamide gel electrophoresis; HEPES, N-2-hydroxyethylpiperazine-N′-2-ethanesulfonic acid; CAPS, 3-[cyclohexylamino]-1-propanesulfonic acid; PVDF, polyvinylidine difluoride; MALDI, matrix-assisted laser desorption ionization; ESIMS-MS, electrospray ionization mass spectrometry–mass spectrometry; SDS, sodium dodecyl sulfate.

are initially expressed by strains of *Clostridium botulinum* as 150-kDa single polypeptides. Posttranslation cleavage by a endogenous trypsin-like protease generates a 50-kDa N-terminal light chain (LC) and a 100-kDa C-terminal heavy chain (HC) that are still connected by a disulfide bond. The LC possesses the toxic, zinc-endopeptidase catalytic domain, but in the absence of HC, it is nontoxic. Recombinant LC expressed by itself is catalytically active (Ahmed and Smith, 2000) and when injected into sea urchin eggs (Ahmed and Smith, 2000; Alderton *et al.,* 2000) or injected into mice after reconstitution of the holotoxin by adding HC (Zhou *et al.,* 1995) is biologically active. The 100-kDa HC can further be proteolyzed into a 50-kDa, N-terminal, membrane-spanning domain (H_n) and a 50-kDa, C-terminal, receptor-binding domain (H_c).

The region connecting BoNT/A LC and HC has seven lysine plus arginine residues (Thompson *et al.,* 1990) that are potential sites of tryptic action. Except for the invariant cysteine at position 430, the last 22-residue sequences of BoNT LCs are also highly divergent. Trypsin or an endogenous protease can cleave at one or more of those seven sites (DasGupta and Dekleva, 1990). This result, along with complete sequence determination of a cyanogen bromide fragment of a BoNT/A LC isolated from the full-length neurotoxin (Krieglstein *et al.,* 1994), indicated proteolytic processing or cleavage at K438 after initial nicking at K448 of BoNT/A holotoxin. This processing, which removes 10 residues from the C-terminus of BoNT/A LC, generating the mature product, is considered to be catalyzed by proteases endogenous to the *Clostridia* producing the toxin (Krieglstein *et al.,* 1994), and in fact a trypsin-like protease was isolated (Dekleva and Dasgupta, 1990). In addition to this C-terminal processing,[6] BoNT/A and BoNT/E LC prepared in 2 M urea from the whole toxin degraded into two large fragments, but the LC in the whole toxin was not affected (DasGupta and Foley, 1989). A recombinant BoNT/A LC expressed in *E. coli* as inclusion bodies also showed protein bands as fragments of the LC, as they reacted in Western blots with antibodies specific for the LC (Ahmed and Smith, 2000). The causes of these degradations in a highly purified native preparation from *Clostridia* (DasGupta and Foley, 1989) and in a recombinant preparation expressed in *E. coli* that lacks any clostridial protease other than the BoNT/A LC (Ahmed and Smith, 2000) remain elusive.

[6] "C-Terminal processing" and "truncation" are interchangeably used in this paper to denote either sequential or random removal of 4–28 residues from the C-terminus of the LC. "Fragmentation" denotes cleavage of the LC or the truncated LC into two large fragments with mass above 15 kDa.

In our continuing efforts to obtain large amount of a soluble, active, and stable BoNT/A LC, we recently purified the protein from soluble fraction of the cell-free extract of *E. coli* expressing the protein. We observed that when stored at 4°C or at 22°C, the purified protein first undergoes a truncation of from 10 to 28 amino acids from the C-terminus followed by fragmentation at peptide bonds in the middle of the molecule. We show here that the protease action is not restricted to lysyl and arginyl bonds. We also demonstrate that the truncation and fragmentation are not due to contaminating protease, but are autocatalytic events occurring at the active site of the BoNT/A LC.

2. MATERIALS AND METHODS

2.1. Chemicals, Buffers, and Reagents

Buffer P (50 mM Na-phosphate, pH 6.5) was used throughout. TPEN and $ZnCl_2$ were from Sigma. Rabbit polyclonal antibodies against a 16-residue N-terminal sequence (PFVNKQFNYKDPVNGV) of BoNT/A LC were produced and affinity purified by Research Genetics (Huntsville, AL). Affinity-purified, peroxidase-coupled goat anti-rabbit and anti-mouse IgG (H + L) and ABTS substrate were from Kirkegaard Perry Laboratories (Gaithersburg, MD). The inhibitor peptide (Ac-CRATKML-NH_2) (Schmidt *et al.,* 1998) was synthesized and purified by Cell Essentials (Boston, MA).

2.2. BoNT/A LC Purification

The rBoNT/A LC was expressed by low-temperature IPTG induction in *E. coli* BL21 (DE3) cells as a soluble protein from a synthetic gene in a pET24a-derived multicopy plasmid (Clontech, Inc.). Construction of the gene and expression of the protein as described (Ahmed and Smith, 2000) was modified as follows: a stop codon replaced the histidine tag at the carboxy terminus of the gene, and induction and expression was at 18°C for 22–24 hr. The LC was purified to near homogeneity by NaCl gradient elution from each of two successive cation exchange columns (MonoS) in buffer P. Details of the expression and purification will be published elsewhere (manuscript in preparation). A typical preparation had a specific activity of 2–3 μmol/min/mg in cleaving the 17-residue substrate peptide (see later) when assayed in the presence of 0.25 mM $ZnCl_2$ (not shown); in the absence of added zinc, activity was 50%. The purified LC was thus partially resolved of the bound zinc. The puri-

fied protein (1–4 ml) in buffer P was stored at −20°C. Under this condition, the protein remains stable and retains its catalytic activity for at least 1 year.

2.3. Proteolysis Experiments

Before each experiment, aliquots of the protein were thawed to room temperature and were immediately passed through a PD-10 column to remove the EDTA. The protein was collected in buffer P and stored on ice. The EDTA-free BoNT/A LC was mixed with predetermined concentrations of $ZnCl_2$, EDTA, TPEN, or the inhibitor peptide (see later), and 20–50 μl was distributed in screw-capped Eppendorf tubes. The tubes were incubated at 4°C or at 22°C. The final concentration of the protein was 0.18–0.20 mg/ml in these incubation mixtures. At various time intervals an equal volume (20–50 μl) of SDS-load buffer was added to a tube for SDS–PAGE analysis.

A 100 mM stock solution of TPEN was prepared in ethanol (95%). Stock solutions of the competitive inhibitor peptide Ac-CRATKML-NH_2 (Schmidt *et al.*, 1998) (5 mM), $ZnCl_2$ (1–4 mM), and EDTA (20 mM) were prepared in buffer P. Unless otherwise mentioned, final concentrations of these reagents in the incubation mixtures with the LC were TPEN 5 mM, EDTA 5 mM, peptide 1 mM, and $ZnCl_2$ 0.25 mM.

2.4. SDS–PAGE, Transfer on PVDF Membrane, and Western Blot

SDS–PAGE under reducing conditions (Laemmli, 1970) was carried out on 1-mm-thick 10% tricine gels (Novex) as described (Schagger and von Jagow, 1987). Samples were prepared in 0.4% SDS, 5% β-mercaptoethanol, 12% glycerol, and 450 mM Tris–HCl, pH 8.45, by boiling for 5 min. The running buffer contained 0.1% SDS in 0.1 M Tris–0.1 M Tricine, pH 8.3. The gels were stained with Coomassie Brilliant Blue. Electrophoretic transfer of peptides from SDS–PAGE gels onto PVDF membrane used 10 mM CAPS–NaOH buffer, pH 11.0, containing 10% methanol as the transfer buffer. Protein bands on the PVDF membranes were visualized by 1 min of staining with Coomassie Brilliant Blue followed by destaining in 10% acetic acid–5% methanol. The stained bands were cut out from the dried membranes for amino-terminal sequence determination. Western blots on nitrocellulose membranes were prepared using a primary polyclonal antibody against a 16-residue N-terminal sequence of BoNT/A LC and a peroxidase-coupled goat anti-rabbit IgG (H + L) as the secondary antibody (Ahmed and Smith, 2000).

2.5. Separation of Peptides with HPLC and Their Characterization by ESIMS-MS

For mass and sequence determination, peptides were separated on an Agilent Technologies Series 1100 liquid chromatograph with a 0.8 × 100 mm Poros-2 R/H column (PerSeptive Biosystems, Inc.). The mobile phase was 0.1% formic acid (solvent A) and 80% acetonitrile in 0.1% formic acid (solvent B). The peptides were eluted with a linear gradient of 0–100% B over 15 min at a flow rate of 0.2 ml/min. The injection volume was 10 μl. The peptides were detected and structurally characterized on a Finnigan LCQ Deca mass spectrometer employing data-dependent MS/MS. Molecular mass was also determined by MALDI-MS with a PE Biosystems Voyager DE instrument. Sinapinic acid was used as the matrix, and the sample was spotted on a stainless steel plate that was not washed with water or TFA. Other conditions in the experiment were accelerating voltage 25,000 V, guide wire voltage 0.3%, and laser 2500.

2.6. Other Analytical Methods

The enzymatic assay was based on HPLC separation and measurement of the nicked products from a 17-residue C-terminal peptide of SNAP-25 corresponding to residues 187–203 (Schmidt and Bostian, 1995). Initially protein concentrations were determined by BCA assay (Pierce) with bovine serum albumin (BSA) as a standard. After it was established by repeated measurements that a 1-mg/ml BoNT/A LC thus determined has $A^{0.1\%}$ (1 cm light path) value of 1.0 at 278 nm (0.98 at 280 nm), protein concentration was determined from absorbance at 278 nm. For comparison, the calculated $A^{0.1\%}$ value of the LC at 280 nm in water (Pace *et al.*, 1995) is 0.948. Absorption spectra were recorded in a Hewlett-Packard 8452 diode array spectrophotometer. The N-terminal amino acid sequence of the LC was determined by Edman degradation in an Applied Biosystems Procise Sequencer in the 0- to 20-pmol detection range.

3. RESULTS

3.1. Cleavage and Fragmentation of BoNT/A LC

Figure 1 shows that the BoNT/A LC undergoes cleavage and fragmentation that increases with time. The intensity of the band representing the full-length LC with a polypeptide mass of ~52 kDa (IA) gradually diminished with time and a new protein band of

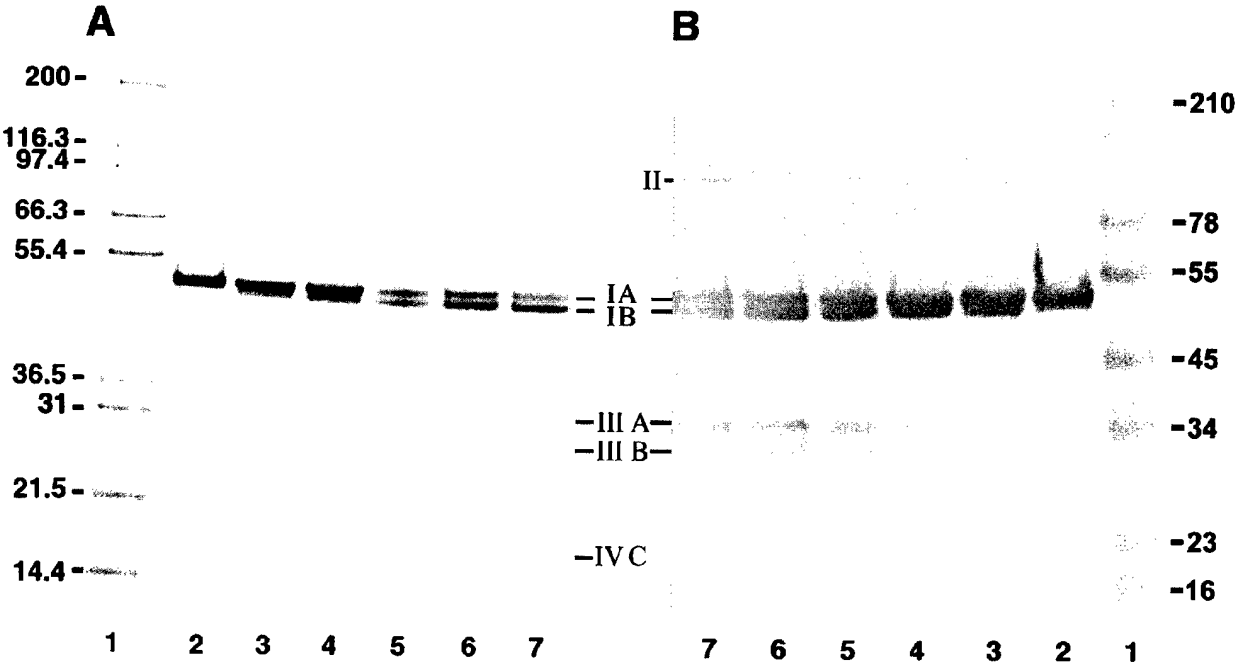


Fig. 1. Time course of proteolysis of BoNT/A LC as followed by SDS–PAGE (**A**) and Western blot (**B**). Aliquots of 25 ml of the LC (0.2 mg/ml) were incubated at 4°C. At intervals (see below), 25 μl of 2× SDS-load buffer was added to an aliquot and boiled. Two SDS gels were run in parallel. One gel was stained by Coomassie (**A**) and the proteins from the other were transferred to a nitrocellulose membrane for Western blot (**B**). Lane 1 in panel A shows Novex Mark-12 molecular weight markers and lane 1 in panel **B** shows the Novex prestained SeeBlue molecular weight markers. In both panels **A** and **B**, lanes 2–7 show 0, 2, 4, 14, 21, and 28 days of incubation, respectively, of LC. Identity of the protein bands between panels **A** and **B** is arbitrary, and the same nomenclature is used throughout the paper.

~50 kDa (IB) appeared in its place. The results suggest truncation of about 2 kDa mass from the full-length LC. In Western blots (Fig. 1B), both IA and IB also reacted with a rabbit polyclonal antibody raised against a 16-residue amino-terminal sequence of LC. This result suggests that the truncation from the full-length LC must occur at the C-terminus. Indeed, amino-terminal sequencing of the isolated, truncated protein (see later) showed the amino terminus was intact. Interestingly, preservation of the N-terminus of full-length BoNT/A neurotoxin was also observed after its posttranslation modification in bacterial culture (DasGupta and Dekleva, 1990). As the truncated protein IB accumulated, a protein band of ~100 kDa (II) appeared that was detected easily in the Western blot (Fig. 1B). Figure 1 also shows that at 2 weeks of incubation, the LC fragmented into IIIA + IIIB and IVC. The larger fragment (IIIA) above the 34-kDa marker was followed by a fainter fragment (IIIB) just below the 34-kDa marker. The results of this time course experiment also suggested that IIIB was formed from IIIA. Both of these fragments must represent the N-terminus of the LC, as they reacted with the antibody (Fig. 1B). On the other hand, a much smaller fragment (IVC) moving faster than the 23-kDa marker was probably the C-terminal fragment, as it failed to react with the antibody (specific for the N-terminus of the LC) in the Western blot. The truncation and fragmentation shown in Fig. 1 were independent of the batch of *E. coli* cell culture or the batch of purification of the LC (data not shown).

3.2. Zinc Accelerates the Truncation and Fragmentation

The BoNT/A LC is known to be highly substrate specific. Therefore, the truncation of about 2 kDa from the C-terminus or fragmentation into larger fragments upon storage of the LC at 4°C described in Fig. 1 might appear to be due to the presence of some contaminating protease in the LC preparation. However, we failed to detect any additional Coomassie-stained protein band when 0.4–4.0 μg of the LC was electrophoresed in the presence of SDS. BoNT/A LC is a zinc-endopeptidase. Figure 2 shows that when LC was incubated with 0.25 mM $ZnCl_2$, the rate of fragmentation was greatly in-

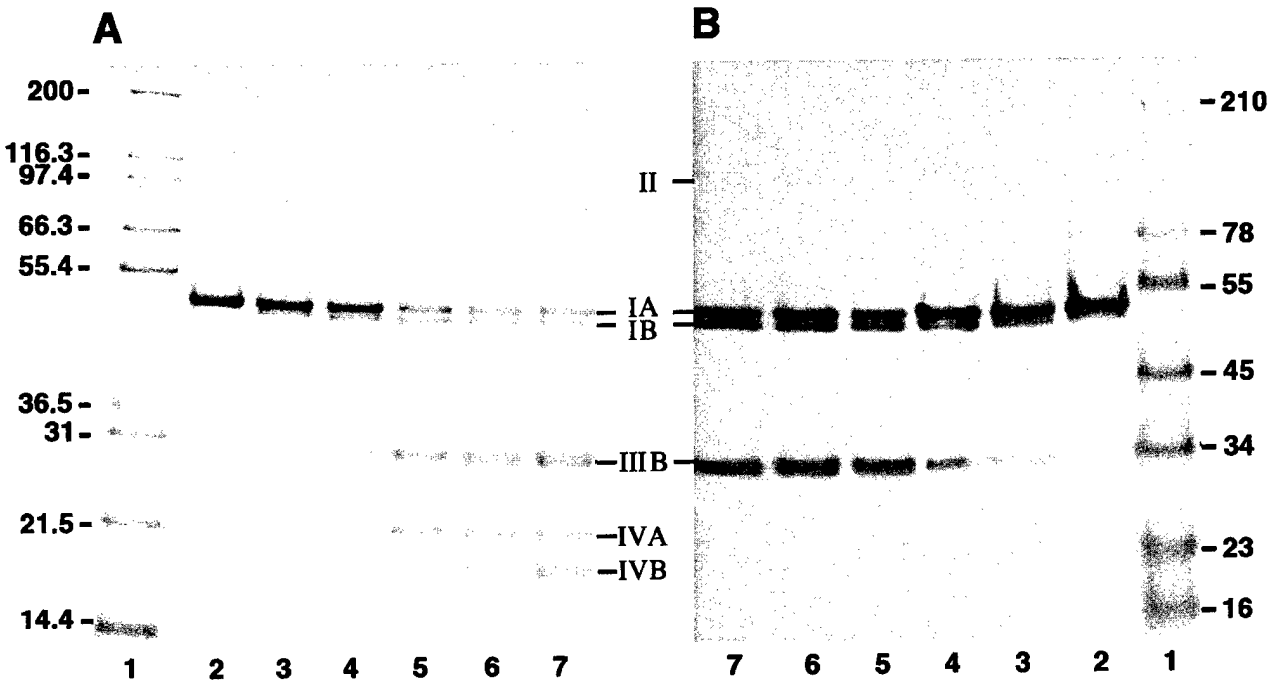


Fig. 2. Enhancement of the proteolysis of BoNT/A LC by $ZnCl_2$ as followed by SDS–PAGE **(A)** and Western blot **(B)**. All conditions are same as in Fig. 1, except that 0.25 mM $ZnCl_2$ was added to the incubation mixture of the LC.

creased so that the antibody-reacting fragment IIIB and an antibody-nonreacting fragment IVA appeared within 2 days of incubation (Fig. 2A, B). Fragment IVB appeared later in the time course. Qualitatively, the results are similar to those in Fig. 1 except that in the presence of $ZnCl_2$, the rate of fragmentation was higher, fragment IIIB was formed without showing the initial formation of IIIA, and initial formation of IVA gave rise to IVB. The rate enhancement by zinc could be partly due to formation of holo-LC from the partially Zn-resolved LC (see Section 2). Because there was no fragment IVC (Fig. 1) detected in this experiment (Fig. 2), zinc must also have a structural role on the LC. From the results shown in Fig. 2A it is not possible to judge if the C-terminal truncation of IA in forming IB and dimerization in forming II precede the fragmentation into III and IV. However, in some other experiments (not shown) using a lower concentration of $ZnCl_2$, we were able to show that formation of IIIB occurred before formation of IB and that fragmentation was the last event.

The rates of C-terminal truncation and fragmentation of LC either in the absence or in the presence of $ZnCl_2$ were much higher when incubated at 22°C than at 4°C (data not shown). In fact, amino-terminal sequence was determined on the fragments generated by incubation at 22°C for 2 days only (see later).

3.3. Amino Acid Sequence of the Small Peptides Generated by C-Terminal Processing

To map the sites of proteolysis, we isolated the small peptides by ultrafiltration of a C-terminally truncated LC mixture. Amino acid sequences of these peptides were determined by ESIMS-MS (Table I). The peptides with G433 at the amino terminus (peptide 4) and K438 at the carboxy terminus (peptide 5) indicated cleavage by a trypsin-like protease on the R432–G433 and K438–T439 bonds, respectively. Of these, only the lysyl bond at K438 was reported to be cleaved by a clostridial endogenous protease or by trypsin (DasGupta and Dekleva, 1990). However, we could not detect a cleavage at the K444–G445 bond as reported before by an endogenous clostridial protease (DasGupta and Dekleva, 1990), nor could we detect cleavage at K440–S441 or at K427–L428 bonds, the other potential sites of tryptic cleavage. Although these results indicated that the LC preparations did not contain a protease activity that could cleave at K427–L428, K440–S441, and K444–G445, it is equally possible that some of the small peptides generated by cleavage at these sites were lost during sample preparation. Interesting findings of this experiment (Table I) are the peptides with N-terminus of T420 (peptide 1) and V431 (peptide 3), as the preceding

Table I. Sequence of the C-Terminal Peptides Generated after Initial Cleavage of the BoNT/A LC[a]

Peptide	Mass[b]	420 425 430 435 440 445 K-N-F-T-G-L-F-E-F-Y-K-L-L-C-V-R-G-I-I-T-S-K-T-K-S-L-D-K-G-Y-N-K
1	2188 (2188)	T-G-L-F-E-F-Y-K-L-L-C-V-R-G-I-I-T-S-K
2	2124 (2112)[c]	[c]C-V-R-G-I-I-T-S-K-T-K-S-L-D-K-G-Y-N-K
3	2008 (2008)	V-R-G-I-I-T-S-K-T-K-S-L-D-K-G-Y-N-K
4	1753 (1753)	G-I-I-T-S-K-T-K-S-L-D-K-G-Y-N-K
5	989 (977)[c]	[c]C-V-R-G-I-I-T-S-K

[a] The peptides were generated by incubating 0.4 mg of the LC in 0.5 ml of buffer P at 4°C for 2 weeks. They were isolated by ultrafiltration through a Centricon CM10 (Amicon) membrane that was previously treated with 10 mM EDTA. The filtrate containing the peptides was stored at −20°C for 1 week before mass and sequence determinations by ESIMS-MS. The sequence on the first row with the numbers above it represents the known C-terminal sequence of the LC (Ahmed and Smith, 2000).

[b] Experimentally determined mass from ESI-MS; calculated mass for the sequence shown is given in parentheses.

[c] The calculated mass was 12.1 Da smaller than the experimental value. Except for cysteine in peptides 2 and 5, the experimentally determined masses of all other amino acid residues agree well with their calculated values. Note that cysteine in peptides 2 and 5 occurred at the N-terminus, but when it was in the middle of the peptide, there was no ambiguity in the results.

residues at F419–T420 and C430–V431 bonds, respectively, are certainly not the sites of "tryptic" cleavage.

The sequence data (not shown) from the ESIMS-MS results for the peptides 2 and 5 agree very well with the residue stretches V432–K449 and with the residue stretches V431–K438, respectively. However the experimentally determined mass for "C430," the residue at the amino side of V431 in both peptides, was greater by 12.1 Dalton than the theoretical mass for cysteine. At this stage, we are uncertain about the discrepancy in the mass of this "cysteine." Chemical modification experiments (data not shown) using iodoacetamide or acidified methanol failed to shift the masses of these peptides, indicating that the suspected "cysteine" did not have a free sulfhydryl group nor was a contaminating aspartic acid. Cysteine in proteins are known to occur as derivatives such as cysteine sulfenic acids (Ahmed and Claiborne, 1992; Claiborne *et al.*, 1999). We are currently attempting to decipher the chemical nature of this "cysteine." If indeed it was a modified C430, cleavages at the carboxy ends of F419, C430, and V431 in addition to R432, K438, and K438 indicate that the proteolytic activity in our preparation was not "tryptic" in nature, but had a broad specificity.

3.4. Identity of the Large Peptides Generated by Fragmentation

The large peptides generated by fragmentation in the middle of the LC were identified by comparing the mass determined by MS with a calculated mass for a stretch of sequence based on the amino-terminal sequence determination (Table II). Agreements between the experimental and calculated values were within 0.07%. Identity of IIIA as having a sequence range of V1–F266 was based on the kinetics of its (and of IVC's) appearance on SDS–PAGE (Figs. 1 and 2) and N-terminal sequence of IVC. The sequence data along with Western blot results clearly demonstrated that the amino terminus of the LC (IA and IB) remained unchanged during the prolonged incubation period. Although we did not determine the C-terminal sequence of the peptides IIIA and IIIB, N-terminal sequences of the peptides IVA, IVB, and IVC (Table II) indicate that fragmentation of IA and IB (Figs. 1 and 2) occurred by cleavage at the Y250–Y251 and F266–G267 bonds. Again, if the cleavages of these tyrosyl and phenylalanyl bonds were catalyzed by a protease, it must have been "nontryptic" in nature. Identity of the peptides IVB and IVC as having F423 at the C-terminal indicated that a C-terminal processing of the LC at F423–E424 remained undetected in our small peptide isolation experiment (see previous section). This result nonetheless supports that C-terminal processing occurred at phenylalanyl bonds in addition to lysyl, arginyl, valyl, and (most likely) cysteinyl bonds.

3.5. Metal Chelator TPEN Inhibits Truncation and Fragmentation

As shown in Fig. 2, if the C-terminal truncation and fragmentation of the LC was indeed dependent on the presence of zinc, removing zinc from the incubation mixture and from the active site of the LC would be expected to abolish the truncation and fragmentation events. However, zinc is very tightly bound to the active

Table II. Identity of the Polypeptides Generated by Proteolysis of the BoNT/A LC

Peptide[a]	Mass (Exp)	Mass (Calc)	Sequence range	N-terminal sequence
IA	51,315	51,318	V1–K448	2-VQFVNKQ
IB	48,866	48,870	V1–Y426	2-VQFVNKQ
II	97,727[b]			
	97,870[b]			
IIIA	n.d.[c]	32,270	V1–F266	2-VQFVNKQ
IIIB	28,111	28,130	V1–Y251	2-VQFVNKQ
IVA	23,207	23,207	Y252–K448	252-YEMSGLE
IVB	20,319	20,319	Y252–F423	252-YEMSGLE
IVC	18,400	18,400	G267–F423	267-GGHDAKF

[a] Peptide designations are from Figs. 1 and 2. Mass was determined by ESIMS-MS. Masses of the peptides IA and IB were determined separately. Peptides were generated by incubating the LC (1.8 mg/ml buffer P) alone or in the presence of 0.25 mM $ZnCl_2$ for 2 days at 22°C. Partial precipitation of the protein was visible after 1 day and was removed by centrifugation before ESI analysis. Masses of IIIB, IVA, and IVB were determined in samples containing $ZnCl_2$, and those of IA, IB, IIIA, and IVC were determined in samples with no $ZnCl_2$. Calculated masses are for the sequence ranges shown based on N-terminal sequence and mass data. The N-terminal sequences were determined separately for IA, IB, and IIIA in solutions and for IIIB, IVA, IVB, and IVC on PVDF membrane after separation by SDS–PAGE and transfer on membrane.

[b] Data from MALDI-MS determined in a sample containing IB with an initial concentration of 0.2 mg/ml.

[c] Mass could not be detected in several experiments, probably due either to precipitation or to irreversible binding to column resin. Although a peptide with a lower mass can have slower mobility than a homologous higher mass peptide in SDS–PAGE due to charge differences (Ahmed *et al.*, 1986), the kinetics of appearance of IIIB from IIIA (Fig. 1) and their identification by N-terminal sequence determination suggest that IIIA must be larger than IIIB. Identity of IIIA as having a sequence of V1–F266 with a mass of 32,270 was based on N-terminal amino acid sequence determination and SDS–PAGE results (Figs. 1 and 2).

site of LC. Extensive treatment with 10 mM EDTA in the cold (Ahmed and Smith, 2000) or with 10 mM EDTA at room temperature (Li and Singh, 2000) failed to completely remove zinc from the active site of the LC. In agreement with these observations, including 10 mM EDTA failed to protect the LC from C-terminal truncation and processing (Fig. 3A). In contrast, the metal chelator TPEN largely protected the LC from truncation and fragmentation (Fig. 3A). We also found that at 1 mM TPEN concentration, the LC showed no activity when assayed for 5 min (data not shown). Because the incubation mixture with TPEN did not contain any exogenous metal or zinc, any chelation by TPEN must have involved the active-site zinc of the LC. Our results also suggest that truncation and fragmentation of the LC upon storage at 4°C or at room temperature were autocatalytic.

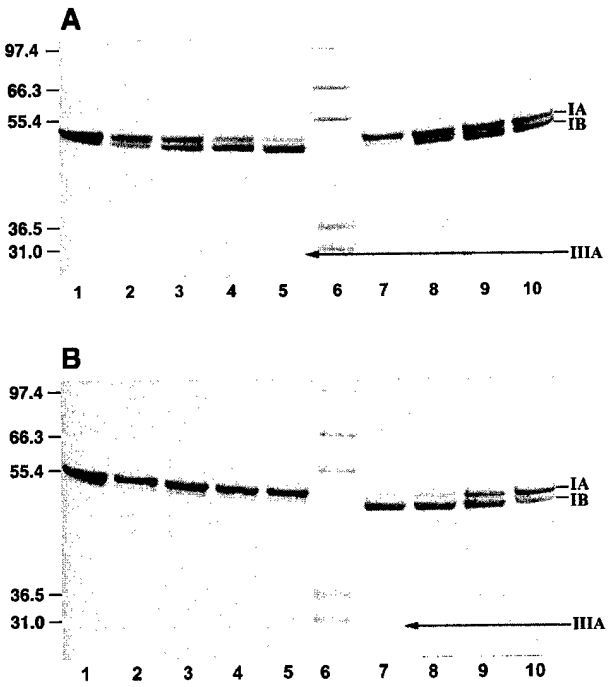


Fig. 3. Protection of BoNT/A LC from proteolysis by the metal chelator TPEN **(A)** and the competitive peptide inhibitor CRATKML **(B)**, followed as a time course by SDS–PAGE. **(A)** The LC (0.2 mg/ml) was incubated in small aliquots with 10 mM EDTA (lanes 2–5) or with 5 mM TPEN (lanes 7–10). Lanes 2 and 7, 3 and 8, 4 and 9, and 5 and 10 show 6, 14, 21, and 28 days of incubation, respectively. **(B)** The LC was incubated with1 mM peptide inhibitor containing 5 mM DTT (lanes 2–5) or without the peptide inhibitor (lanes 10–7) at 4°C. DTT, which does not have an effect on proteolysis (not shown), was added to maintain the peptide in monomer form. Lanes 2 and 10, 3 and 9, 4 and 8, and 5 and 7 show 6, 14, 21, and 28 days of incubation, respectively. In both panels **A** and **B**, lane 1 represents LC alone at day 0, and lane 6 has molecular weight markers (labels on left). The protein band IIIA (see Fig. 1) was faint in this experiment and was not captured in the photographic reproduction; therefore its location in the original gel is shown by arrows in the figure. Note that (a) presence (lanes 2–5, A) and absence (lanes 10–7, B) of EDTA had little effect on proteolysis of IA to IB and finally to IIIA, (b) TPEN (lanes 7–10, A) significantly reduced the rate of conversion of IA to IB and prevented formation of IIIA during the course of the experiment, and (c) the peptide inhibitor (lanes 2–5, B) drastically reduced the proteolysis of IA to IB and prevented the formation of IIIA.

3.6. A Specific Competitive Inhibitor of LC Activity Was an Effective Inhibitor of Truncation and Fragmentation

Autocatalytic truncation and fragmentation of proteins can arise from chemical catalysis and from enzymatic catalysis. To differentiate these two possibilities, we used a peptide specifically synthesized as a competitive inhibitor of BoNT/A proteolytic activity (Schmidt

et al., 1998). This peptide inhibitor, with a sequence of CRATKML, competitively inhibits the cleavage of a 17-residue substrate peptide based on SNAP-25 by BoNT/A neurotoxin with a K_i of 2 μM (Schmidt *et al.,* 1998). We found that at 1 mM inhibitor peptide concentration, the LC showed no activity when assayed for 5 min (data not shown). Figure 3B shows that when the LC was incubated with 1 mM peptide inhibitor, both C-terminal truncation and fragmentation at the interior of LC were largely prevented. In the presence of the peptide inhibitor, however, the LC underwent a very slow cleavage, as can be expected in an enzymatic activity with a competitive inhibitor. Densitometric scanning of the gel showed that after 28 days, in the presence of the peptide inhibitor, less than 10% of the LC (IA) was converted into the C-terminally truncated form (IB). In contrast, in the absence of the peptide inhibitor, more than 80% of the LC (IA) was converted into the truncated form (IB). Results of this experiment prove that loss of 10–28 residues from the C-terminus of LC followed by fragmentation into two major peptides (Figs. 1 and 2, Tables I and II) occurred at the active site of the LC and that these reactions were enzymatic. The results also provide direct evidence that the cleavage reactions were not due to any contaminating protease in the preparation of the LC.

4. DISCUSSION

4.1. C-Terminal Processing and Fragmentation of the LC Are Autocatalytic

Our studies provide the first evidence in favor of autocatalysis of BoNT/A LC. Several lines of evidence also indicate that this autocatalysis is enzymatic in nature:

1. Highly purified LC (Fig. 1) underwent proteolysis (Fig. 1) regardless of its batch of purification and batch of *E. coli* cell growth expressing the protein.
2. Zinc, a cofactor bound at the active site of BoNT/A LC (Fu et al., 1998), whose addition (as $ZnCl_2$) to the reaction mixture accelerates its enzymatic catalysis (Ahmed and Smith, 2000), also enhances BoNT/A proteolysis (Fig. 2).
3. The heavy metal chelator TPEN that inhibits the enzymatic activity of LC also inhibited the proteolysis of LC (Fig. 3A). TPEN is the most effective metal chelator antagonist in the BoNT-induced blockade of transmission in mouse nerve-hemidiaphragm preparations (Simpson et al., 1993; Sheridan et al., 1997) that inhibits the enzymatic activity of LC (this study) and protects animals against BoNT/A poisoning (Adler et al., 1997).
4. The heptapeptide CRATKML that was shown to be a competitive inhibitor of BoNT/A catalytic (proteolysis) activity (Schmidt et al., 1998) and that inhibits LC catalytic activity (this study) effectively protects the LC from proteolysis (Fig. 3B).

Cleavage of highly purified LC from various batches first reduced the possibility that the proteolysis of LC was due to a contaminating protease. By enhancing the proteolysis with zinc and inhibiting it with TPEN (a metal chelator), we suggest that the proteolysis was catalyzed by the LC itself. Finally, protection from proteolysis by the competitive inhibitor validated that the proteolysis was enzymatic and autocatalytic. An earlier study on LC prepared in 2 M urea from the whole toxin purified from *Clostridium* reported that the Y250–Y251 bond was cleaved over time (DasGupta and Foley, 1989). From the facts that the protein was purified in the presence of aprotinin and PMSF, that no fragmentation was observed in the full-length toxin, and that no other Y–Y bond in the molecule was cleaved, presence of a contaminating protease in the LC preparation was considered unlikely. If cleavage was due to a contaminating protease, then proteases having this same bond specificity must be common to *E. coli* as well as *Clostridium,* and must have physical properties that cause them to copurify with the recombinant LC as well as whole toxin; this is not very likely. Instead of enzymatic catalysis, an alternate explanation of a sequence-specific chemical cleavage was sought (DasGupta and Foley, 1989). We previously also observed protein bands recognized by LC-specific antibody that appeared to arise by fragmentation of the LC expressed as inclusion bodies in *E. coli* (Ahmed and Smith, 2000). This study also found fragmentation by cleavage at the Y250–Y251 bond and an additional fragmentation at F266–G267. C-terminal processing and fragmentation with an LC expressed and purified in soluble form from *E. coli* avoiding the use of the denaturants like urea (DasGupta and Foley, 1989) or sarkosyl (Ahmed and Smith, 2000) eliminate the possibility of exposing a cleavage site by prior use of a denaturant. An autocatalytic enzymatic action of the LC is the most logical explanation. Direct evidence in support of the autocatalytic fragmentation comes from a recent preliminary report in which the Y250–Y251 bond was found at the active site of a truncated LC and as cleaved (Knapp *et al.,* 2000) (see later).

4.2. Fragmentation Follows a Stable Dimer Formation

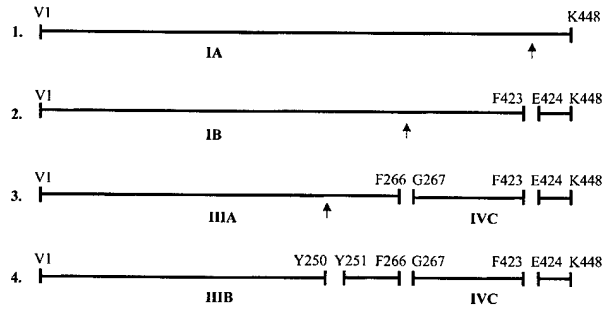
During purification from inclusion bodies, we observed a protein band in reducing SDS–PAGE that appeared to be a dimer of the LC, in addition to several smaller fragments (Ahmed and Smith, 2000). In the experiments shown in Fig. 1, we observed the gradual accumulation of a species (II) with a mass of 97,870 (Table II) that appeared to follow the accumulation of the C-terminally truncated LC (IB) with a mass of 48,892 (Table II). Species II thus was a dimer arising from species IB. Presence of a non-covalently bound dimer in a reducing SDS–PAGE experiment is generally unlikely but not uncommon (Klatt *et al.,* 1995; Ettinger *et al.,* 2000). Consistent time-dependent appearance of this species in several experiments led us to conclude that an SDS-stable dimer was formed after the removal of 10–28 amino acid residues from the C-terminus of LC. In a dimer, the Y250–Y251 bond of one IB chain may be juxtaposed to the active site of the complementary chain for cleavage. This explanation is consistent with the kinetics of formation of IB, II, and IIIA and IVC (Fig. 2 and Scheme II). In the three-dimensional structure of the LC determined at 3.3 Å resolution for the whole BoNT/A toxin (Lacy *et al.,* 1998; Lacy and Stevens, 1999), residues 432–448 at the C-terminus show no electron density. Residues 420–431 containing one β-strand (β15) appear to be more flexible than the rest of the molecule. Residues 233–259 containing the Y250–Y251 bond formed a loop (loop C) bordering the active-site gorge, and the F266–G267 bond occurs just after the helix 4 (residues 260–266) ends. A visual inspection of the structure suggests that proteolytic removal of the C-terminus should eliminate the steric hindrance for the loop C to be accommodated in the active site that is situated in a 35-Å-deep pit (Lacy *et al.,* 1998; Lacy and Stevens, 1999) in forming the dimer. Direct support for the formation of a stable dimer comes from a recent preliminary report on the structure of a C- and N-terminal truncated LC (Kadkhodayan *et al.,* 2000) determined by X-ray crystallography (Knapp *et al.,* 2000). The molecule formed a dimer in the asymmetric unit in which the Y250–Y251 bond was cleaved, near the active site, and the resulting carboxylate group remained coordinated to the catalytic zinc (Knapp *et al.,* 2000). Two related reports on factors inducing SDS-stable dimer formation that may involve a carboxylate come from addition of tetrehydrobiopterin and arginine to nitric oxide synthase (Klatt *et al.,* 1995) and the presence of a specific glutamic acid residue in a class II MHC allele (Ettinger *et al.,* 2000). Our results of proteolysis in the presence of $ZnCl_2$ (Fig. 2 and Scheme II), on the other hand, also showed that removal of the C-terminus could not be a prerequisite for a dimer formation, as fragmentation occurred before accumulation of IB. This indicates a structural role for zinc.

4.3. A Structural Role for Zinc

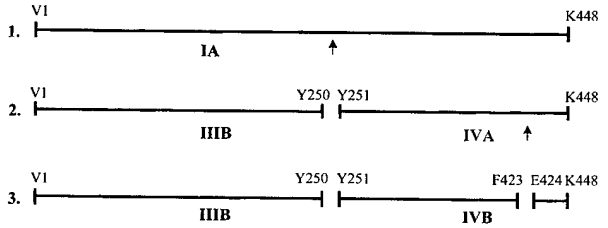
Zinc, the natural cofactor of the LC, is tightly bound at its active site. Most preparations of the LC that were purified through two columns and a dialysis step in the presence of EDTA retained at least 50% of its activity in the absence of added zinc (see Section 2). In the experiments described here, adding zinc increased the rate of fragmentation (compare Fig. 2 with Fig. 1) partly due to formation of holo-LC from the partially Zn-resolved LC. Apart from these measurable differences, there was a qualitative difference in the self-proteolysis in the presence and absence of added $ZnCl_2$ that are summarized in Schemes I and II, respectively. In the absence of $ZnCl_2$, C-terminal processing occurred followed by fragmentation from initial cleavage at the F266–G267 bond (Scheme I). In the presence of $ZnCl_2$, fragmentation resulted from cleavage at the Y251–Y252 bond bypassing the C-terminal processing (Scheme II). Zinc at the active site or in excess in solution thus can changed the preference of the cleavage site from F266–G267 to Y251–Y252, obviously by some conformational or structural change in the molecule. Zinc thus must have a structural role in addition to its essential catalytic role. In fact, a structural role of zinc in maintaining the tertiary structure of the LC (Li and Singh, 2000) and of the full-length BoNT/A (Fu *et al.,* 1998) was recently reported. Moreover, our failure to observe a conversion of IVB to IVC (1–2, Schemes I–II) may indicate a changed conformation of the N-terminus of IVB from that which occurs in the middle of LC (IA or IB). Indeed, a specific structural element has been proposed to be a requirement for the true substrates of BoNT neurotoxins (Washbourne *et al.,* 1997; Lebeda and Olson, 1994; Rossetto *et al.,* 1994) rather than its primary structure alone. This analogy provides additional support in favor of an autocatalysis of the LC.

4.4. Specificity of the Proteolysis

BoNT/A and its LC are specific for only the Q197–R198 bond of SNAP-25. By demonstrating the autocatalytic action of LC in this report, we found that



Scheme I. Steps in the self-proteolysis of BoNT/A LC in the absence of added zinc. Arrows show the sites of proteolysis. Full-length LC is denoted by IA. The fragments 1B, IIIB, and IVC correspond to the fragment designations in Fig. 1. The primary event is the C-terminal truncation to form IB followed by cleavage between Y286 and G287 producing IIIA and IVC. The fragment IIIA in turn is further proteolyzed between Y251 and Y252 to generate IIIB. Lengths of the fragments (e.g., 1V–K448) are based on mass determined by MALDI-MS and N-terminal amino acid sequence shown in Table II. The C-terminal peptide E424–K448, although shown here as a single peptide for convenience, is in fact is a mixture of several peptides (see Tables I and II).



Scheme II. Steps in the self-proteolysis of BoNT/A LC in the presence of added zinc. Arrows show the sites of proteolysis. The fragments IIIB, IVA, and IVB correspond to the fragment designations in Fig. 2. Unlike the steps shown in Scheme I, IA may bypass the C-terminal truncation and initial formation of IIIA but undergo proteolysis between Y251 and Y252 in directly forming IIIB. The fragment IVA is further cleaved into IVB. Although a C-terminal cleavage of IVB into IVC is possible, we did not observe (see Fig. 2) this species in the presence of added zinc. See Fig. 2 and Scheme I for other explanations.

Y250–Y251, F266–G267 at the middle of the molecule, and F419–T420, F423–E424, R432–G433, K438–T439, C430–V431, and L429–C430 (most likely) bonds at the end of the molecule were also recognized as sites of proteolytic action. This broad specificity of the cleavable bonds is not surprising if we consider that substitutions at P2, P1, and P2′ sites of a 17-residue peptide did not abolish its substrate property for BoNT/A (Schmidt and Bostian, 1997). Because not all Y–Y, F–E, F–G, R–G, and K–T bonds of the light chain were cleaved, it is apparent that the cleavable bonds became accessible to autocatalysis during storage, probably by assuming a favorable structure or conformation. We did not determine the rates of autocatalysis; a rough estimate, however, indicated that the rate of cleavage of the Q197–R198 bond in a 17-residue C-terminal peptide of SNAP-25 by the BoNT/A LC was 10^5 times higher than the autocatalytic cleavage. We are currently investigating the relationship between autocatalysis of LC and its catalytic activity on substrates.

4.5. Conclusions

We provide evidence for the first time that proteolytic C-terminal processing and a proteolytic fragmentation reaction in a soluble BoNT/A LC are enzymatically catalyzed by the LC itself. We found that the specificity of proteolysis of the LC was not restricted to a particular Q–R bond of SNAP-25 but encompassed specific Y–Y, F–G, V–R, R–G, K–T, C–V, and probably L–C bonds in the LC when they occurred in a favorable conformation.

ACKNOWLEDGMENTS

We thank Laura Shanholtz for the proteolysis experiments and enzyme assays, Dr. James J. Schmidt for N-terminal sequence determination, helpful comments, and discussions, and Dr. Frank Lebeda for a critical reading of the manuscript.

REFERENCES

Adler, M., Dinterman, R. E., and Wannemacher, R. W. (1997). *Toxicon* **35**, 1089–1100.
Ahmed, S. A. and Claiborne, A. (1992). *J. Biol. Chem.* **267**, 3832–3840.
Ahmed, S. A. and Smith, L. A. (2000). *J. Protein Chem.* **19**, 475–487.
Ahmed, S. A., Fairwell, T., Dunn, S., Kirschner, K., and Miles, E. W. (1986). *Biochemistry* **25**, 3118–3124.
Alderton, J. M., Ahmed, S. A., Smith, L. A., and Steinhardt, R. A. (2000). *Cell. Calcium* **28**, 161–169.
Claiborne, A., Yeh, J. I., Mallett, T. C., Luba, J., Crane, E. J., 3rd, Charrier, V., and Parsonage, D. (1999). *Biochemistry* **38**, 15407–15416.
Dalbey, R. E. and Kuhn, A. (2000). *Annu. Rev. Cell. Dev. Biol.* **16**, 51–87.
DasGupta, B. R. and Dekleva, M. L. (1990). *Biochimie* **72**, 661–664.
DasGupta, B. R. and Foley, J., Jr. (1989). *Biochimie* **71**, 1193–1200.
Dekleva, M. L. and Dasgupta, B. R. (1990). *J. Bacteriol.* **172**, 2498–2503.
Ettinger, R. A., Liu, A. W., Nepom, G. T., and Kwok, W. W. (2000). *J. Immunol.* **165**, 3232–3238.
Fu, F. N., Lomneth, R. B., Cai, S., and Singh, B. R. (1998). *Biochemistry* **37**, 5267–5278.
Kadkhodayan, S., Knapp, M. S., Schmidt, J. J., Fabes, S. E., Rupp, B., and Balhorn, R. (2000). *Protein Expr. Purif.* **19**, 125–130.
Klatt, P., Schmidt, K., Lehner, D., Glatter, O., Bachinger, H. P., and Mayer, B. (1995). *EMBO J.* **14**, 3687–3695.
Knapp, M., Segelke, B., Balhorn, R., and Rupp, B. (2000). The crystal structure of botulinum toxin A zinc protease domain. Presented at

the 37th Annual Meeting of the Interagency Botulinum Research Coordinating Committee, Alisomar, California.
Krieglstein, K. G., DasGupta, B. R., and Henschen, A. H. (1994). *J. Protein Chem.* **13**, 49–57.
Lacy, D. B. and Stevens, R. C. (1999). *J. Mol. Biol.* **291**, 1091–1104.
Lacy, D. B., Tepp, W., Cohen, A. C., DasGupta, B. R., and Stevens, R. C. (1998). *Nat. Struct. Biol.* **5**, 898–902.
Laemmli, U. K. (1970). *Nature* **227**, 680–685.
Lebeda, F. J. and Olson, M. A. (1994). *Proteins* **20**, 293–300.
Li, L. and Singh, B. R. (2000). *Biochemistry* **39**, 10581–10586.
Montecucco, C. and Schiavo, G. (1995). *Q. Rev. Biophys.* **28**, 423–472.
Pace, C. N., Vajdos, F., Fee, L., Grimsley, G., and Gray, T. (1995). *Protein Sci.* **4**, 2411–2423.
Rossetto, O., Schiavo, G., Montecucco, C., Poulain, B., Deloye, F., Lozzi, L., and Shone, C. C. (1994). *Nature* **372**, 415–416.
Schagger, H. and von Jagow, G. (1987). *Anal. Biochem.* **166**, 368–379.
Schiavo, G., Rossetto, Tonello, F., and Montecucco, C. (1995). Intracellular targets and metalloprotease activity of tetanus and botulinum neurotoxins. In *Clostridial Neurotoxins: The Molecular Pathogenesis of Tetanus and Botulism* (Montecucco, C., ed.), Springer, New York, pp. 257–273.
Schmidt, J. J. and Bostian, K. A. (1995). *J. Protein Chem.* **14**, 703–708.
Schmidt, J. J. and Bostian, K. A. (1997). *J. Protein Chem.* **16**, 19–26.
Schmidt, J. J., Stafford, R. G., and Bostian, K. A. (1998). *FEBS Lett.* **435**, 61–64.
Sheridan, R. E., Deshpande, S. S., Nicholson, J. D., and Adler, M. (1997). *Toxicon* **35**, 1439–1451.
Simpson, L. L., Coffield, J. A., and Bakry, N. (1993). *J. Pharmacol. Exp. Ther.* **267**, 720–727.
Strasser, A., O'Connor, L., and Dixit, V. M. (2000). *Annu. Rev. Biochem.* **69**, 217–245.
Thompson, D. E., Brehm, J. K., Oultram, J. D., Swinfield, T. J., Shone, C. C., Atkinson, T., Melling, J., and Minton, N. P. (1990). *Eur. J. Biochem.* **189**, 73–81.
Washbourne, P., Pellizzari, R., Baldini, G., Wilson, M. C., and Montecucco, C. (1997). *FEBS Lett.* **418**, 1–5.
Zhou, L., de Paiva, A., Liu, D., Aoki, R., and Dolly, J. O. (1995). *Biochemistry* **34**, 15175–15181.